Front. Electr. Electron. Eng. China (2006) 3: 325–329
DOI 10.1007/s11460-006-0044-x

RESEARCH ARTICLE

ZHAO Hai, SU Wei-ji, XU Ye, ZHANG Xin

The effect of Internet separation degree time sensitivity on transmission

Abstract Separation degree is a standard measure for complex network research. Whatever its scale or its increase makes the Internet take on a complex network character. Because of the development of complex network theory and the continuous evolution of the Internet, it is a key problem that uses complex network theory to research the Internet nowadays. In this paper, the Internet separation degree is put forward. The time series stochastic process model of the Internet separation degree is established. According to actual data, the Internet separation degree time sensitivity model (ISDTSM) is established and the effect of time sensitivity of the Internet separation degree to the Internet IP level transmission is computed. Finally the Internet separation and IP transmission during 2008 Beijing Olympic Games were forecasted by using the model.

Keywords separation degree, clustering coefficient, average path length, hops, RTT, complex network, small world phenomenon

1 Introduction

The Internet is a complex network system that has large redundancy. Researching the complex network character of the Internet and discovering its inherent mechanism are inevitable courses and the foundation of further development and use of the Internet. In recent years both the development of complex networks and the development of the Internet have impelled more and more scientists to research the Internet using complex network theory [1, 2].

The beginning of the research of complex networks was the famous mail carrier experiment of Stanley Milgram. The experiment showed that the average distance between sources and targets, if packets were delivered successfully, would be no more than six. This phenomenon was called "Six Degrees of Separation". In 1998, Watts and Strogatz, in their "small-world" theory, discovered that the small world network has larger clustering coefficient and smaller separation degree synchronously. From then on, computer scientists were enlightened to research network by complex network theory. Yook et al. computed Internet' Small World character measurement on the AS level [3]. Adamic and Huberman computed Internet' Small World character measurement on WWW web site level [4].

According to the global data offered by CAIDA's 30 monitors from July 2001 to June 2004, some Internet route information were sampled, computed and analyzed. The concept and algorithm of Internet separation degree, which were based on hops, were put forward. The time sensitivity of the Internet separation degree was analyzed and the Internet separation degree time sensitivity model was deduced. It can use statistics to analyze the relation between the Internet separation degree and transmission. Finally the Internet separation degree and IP route time for the 2008 Beijing Olympic Games were forecasted

2 Complex network and separation degree

2.1 The history of complex network research

A network is a system that includes a great deal of units and reciprocity in these units. In the last two hundred years, there were three phrases that says complex network research experienced: regular network phrase, which was put forward in the eighteenth century; random network phrase, which was put forward in the middle of the ninteenth century; and complex network phrase, which was put forward in modern times.

A regular network is a kind of network whose relation between any nodes could be expressed by a set of regular structure, namely that all the nodes' degrees are the same. Random network is the kind of network in which N nodes and C_N^2 edges, all edges are treated equally and they are

Translated from *Journal on Communications*, 2005, 26(9): 1–8 (in Chinese)

ZHAO Hai(✉), SU Wei-ji, XU Ye, ZHANG Xin
School of Information Science and Engineering,
Northeastern University, Shenyang 110004, China
E-mail : zhaoh@mail.neu.edu.cn

presented with probability p, and not presented with probability $(1-p)$.

Scientists discovered that a lot of actual networks are not only regular networks but also random networks and they have different statistical character from the former two. The network above is called complex network by scientists. Small world network and scale free network are the most important in complex networks.

A network that has a higher clustering coefficient and lower separation degree is called a small world network. This phenomenon is called the small-world phenomenon.

Because of the lack of a scale to describe its character, the network, whose degree distribution has a power-law tail, is called scale free network.

2.2 Separation degree

Separation degree and clustering coefficient are basic scales to research complex networks. According to graph theory, graph G is a two set (V, E), E is subset of $V \times V$. Set V is called vertex set and set E is called edge set. If $\{x, y\} \in E$, there is an edge from x to y in graph G, namely $x \to y$. In the graph, vertex x is called starting vertex and y is called ending vertex. There is an edge $x \to y$ and an edge $y \to x$ synchronously in a graph synchronously. This kind of graph is called an undirected graph. In graph G, $x-y$ is a route P, in which vertex x regards as starting vertex and vertex y regards as ending vertex, is a set composed of a series of edges end-to-end, namely $E(P)=\{x_0x_1, x_1x_2, \cdots, x_{l-1}, x_l\}$, and $x_0 \equiv x$, $x_l \equiv y$, $x_i \neq x_j$, $\forall 0 \le i < j \le l$. The number of edges l is called length of P. If $x_0\ x_l \in E$, then the set of edge $\{x_0\ x_1, x_1\ x_2, \cdots, x_{l-1}\ x_l,\ x_l\ x_0\}$ is a called circle. Its length is $l+1$. In graph G, the length of the shortest $x-y$ route is called the distance of vertex x and y, namely $d_G(x, y)$.

Definition 1 (Separation Degree) Suppose that a normal topology network graph G is a connected graph in which there are N vertexes and k edges that haven't right and $K << N(N-1)/2$, namely G is a sparse graph. The associated matrix of G could be expressed as $A=\{a_{ij}\}$. If there is an edge between i and j, then $a_{ij}=1$, or else $a_{ij}=0$. By using graph theory and network theory, the shortest route matrix could be gotten by $D=\{d_{ij}\}$. The separation degree is

$$L = \frac{1}{N(N-1)} \sum_{i \neq j} l_{ij} \tag{1}$$

L is the mean distance of any two vertexes in graph G, l_{ij} is the factual distance of vertex i and j [2].

The separation degree exhibits the network's connectivity. Lower connectivity shows that the number of middle vertexes, which data packages are passed, is fewer and transmission time in the network is shorter. Higher connectivity shows that the number of middle vertexes, which data packages are passed, is more and transmission time in the network is longer.

3 Internet separation degree

3.1 The state of complex network research in the world

Yook et al. computed the small world character of the Internet on the AS level. Adamic and Huberman the computed the small world character of the Internet on the WWW level, (see Table 1).

Table 1 Small world character of Internet

Network	Size	l	C
Internet [3], domain level	3 015–6 209	3.7–3.76	0.18–0.3
WWW [4], site level	153 127	3.1	0.107 8

In the table, the actual research network object (Network), the size of network (Size), the mean of shortest distance (l) and the mean of clustering coefficient (C) are included. By comparing actual network with a random network whose vertexes and edges are the same, the small world character of the Internet can be discovered [5].

According to scale free network character, Faloutsos et al and Govindan and Tangmunarunkit computed the scale free character of the Internet on the AS level and router level. The results aresummarized in Table 2.

Table 2 Scale free character of Internet

Network	Size	l_{real}	l_{rand}
Internet [6], domain	3 015–4 389		6.3
Internet [6], router	3 888	12.15	8.75
Internet [7], router	150 000	11	12.8

In the table, the actual research network object (Network), the size of network (Size), the mean of shortest distance (l_{real}) and the mean of shortest distance based on random network (l_{rand}) are included. It is obvious that the scale free network has characters that are similar to the character of the small world network [5].

3.2 Internet separation degree algorithm

In TCP/IP protocol, hops illuminates the number of routers from the source node to the destination node in the Internet. Using hops, route efficiency and network transfer ratio of one route in the Internet can be calculated. The complex network character of the Internet, which is ascertained via hops, can reflect actual topology rules of the complex network of the Internet .

Definition 2 (Separation Degree of Internet) is the mean of hops, namely $\overline{\text{HOPS}}$, computed by the formula as follows:

$$\overline{\text{HOPS}} = \frac{1}{N}\sum_{i=1}^{N}(\text{hops})_i \quad (2)$$

in the formula, $(\text{hops})_i$ is *i*th hops a link that is made up of any two IP nodes *A*, *B*. *N* is the total number of links.

3.3 Interent separation degree evolution

It is not practical to count the hops of each route of any two Internet nodes and then calculate $\overline{\text{HOPS}}$ of the Internet. Utilizing 30 global central routers as source nodes, we scanned the Internet and obtained route information. Global Internet data from July 2001 to June 2004, was analyzed. The result is illustrated in Fig. 1. It is obvious that hops of any node have trend, time sensitivity and randomicity.

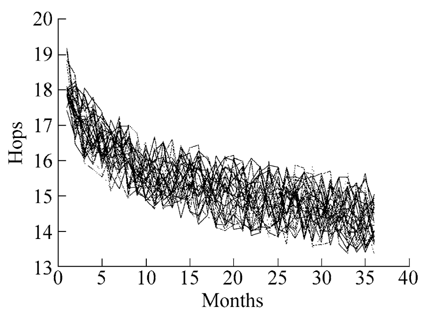

Fig. 1 Hops's trend of 30 nodes of global Internet

In Fig. 1, *Y*-axis is hops; *X*-axis months, range is from 0 to 36, which represents the time's range from July 2001 to June 2004.

If stochastic variable hops, which are from different nodes at different time, are regarded as different tests of stochastic variable $\text{HOPS}(t)$, for a certain time t, $\text{HOPS}(t)$ is a stochastic variable. However when t changes in a certain time range [*a*, *b*], a family stochastic variable can be found and denoted by $\{\text{HOPS}(t), t \in [a,b]\}$, which can be called time sensitivity evolving stochastic process.

For the stochastic process $\{\text{HOPS}(t), t \in [a,b]\}$, its mean function is

$$\overline{\text{hops}_t} = E(\text{HOPS}(t)) \quad (3)$$

$\overline{\text{hops}_t}$ expresses the swing of stochastic process $\text{HOPS}(t)$ in each time. Its trend curve is shown in Fig. 2.

variance function is expressed in the following formula:

$$\text{var}(\text{HOPS}(t)) = E[(\text{HOPS}(t) - \overline{\text{hops}_t})^2] \quad (4)$$

$\text{var}(\text{HOPS}(t))$ expresses departure degree of stochastic process $\text{HOPS}(t)$ to $\overline{\text{hops}_t}$, as Fig. 2 shows.

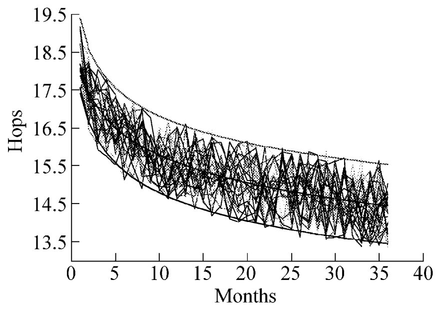

Fig. 2 Curve of mean function and variance function

From July 2001 to June 2004, the values of the mean and variance are listed in Table 3.

Table 3 The value of mean and variance

Time	$\overline{\text{hops}_t}$	var(HOPS(t))
2001.7	18.173 70	0.202 345 0
2001.10	16.732 90	0.359 630 0
2002.2	16.012 50	0.419 666 0
2002.6	15.591 10	0.318 756 0
2002.10	15.292 10	0.424 288 0
2003.2	15.060 20	0.373 680 0
2003.6	14.870 70	0.353 687 0
2003.10	14.710 50	0.379 279 0
2004.2	14.571 80	0.235 789 0
2004.6	14.449 30	0.249 568 0

Via regression analyzsis, the conclusion could be drawn that $\overline{\text{hops}_t}$ conforms with the power-law relation. The function is shown as follows:

$$\overline{\text{hops}_t} = b_0\, t^{b_1} \quad (5)$$

where b_0 = 18.297 9, b_1 = – 0.065 2. The formula $\overline{\text{hops}_t}$ = $18.297\,9t^{-0.0652}$ was called the model of time series stochastic process of the Internet separation degree.

From Figs. 3 and 4, it can be seen that the stochastic process $\text{HOPS}(t)$ is not a calm stochastic process. This is because the first rank calm stochastic process needs a constant mean function namely EX(T) = m. The second rank calm stochastic process need both its mean function and its variance function to be constant, namely var(T) = m. From the figures, we also find that it is not a Poisson stochastic process because it does not conform with the Poisson distribution. Its future state doesn't depend on the current state, so it is not a Markov stochastic process. Therefor the Internet separation degree has its own evolving rules, which cannot be described by traditional stochastic process models.

According to the separation evolving model Eq. (5), it can be estimated that the global mean HOPS is 13.706 9 during the 2008 Beijing Olympic Games.

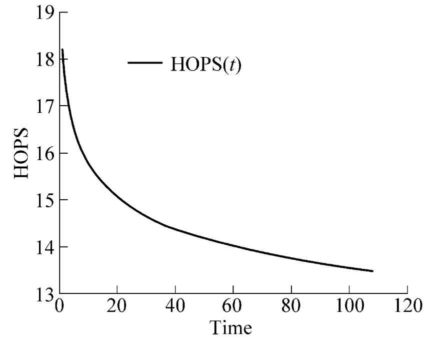

Fig. 3 The curve of mean function

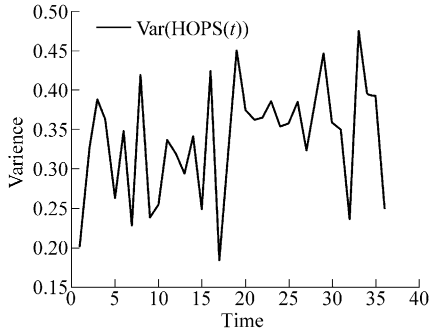

Fig. 4 The curve of variance function

3.4 The effect of internet separation to transmission

In Ref. [2], Watts and Strongatz put forth that the transmission time of all networks, from regular network to random network, conformed with its separation degree. However it does not take into consideration the time scale of the separation degree, i.e., using separation, the transmission time cannot be calculated well. In order to calculate the network transmission time, other physics parameters must be found and a specific transmission model must be created. At present, the most broad and robust models are the SIR model and SIS model [8–10].

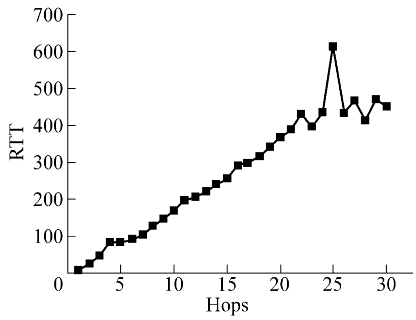

Fig. 5 The correlation analyzing of hops and RTT

Because of the particularity of topology and protocol of the Internet, by correlation analyzsis of hops and RTT, the effect of separation degree to transmission can be determined. The statistic results of 30 global central routers is shown in Fig. 5. In order to constitute the relation between hops and RTT, regression analyzsis between hops and RTT is done. At first, ten curve models are supposed to be as follows:

Liner: $Y = a + b_1 X + b_2 X$ (6)

Quadratic: $Y = b_0 + b_1 X + b_2 X^2$ (7)

Compound: $Y = b_0 \times b_1^{\,X}$ (8)

Growth: $Y = \mathrm{e}^{(b_0 + b_1 X)}$ (9)

Logarithmic: $Y = b_0 + b_1 \ln X$ (10)

Cubic: $Y = b_0 + b_1 X + b_2 X^2 + b_3 X^3$ (11)

S: $Y = \mathrm{e}^{(b_0 + \frac{b_1}{X})}$ (12)

Exponential: $Y = b_0 \, \mathrm{e}^{b_1 X}$ (13)

Inverse: $Y = b_0 + \dfrac{b_1}{X}$ (14)

Power: $Y = b_0 \, X^{b_1}$ (15)

The result of the fitting curve is show in Fig. 6.

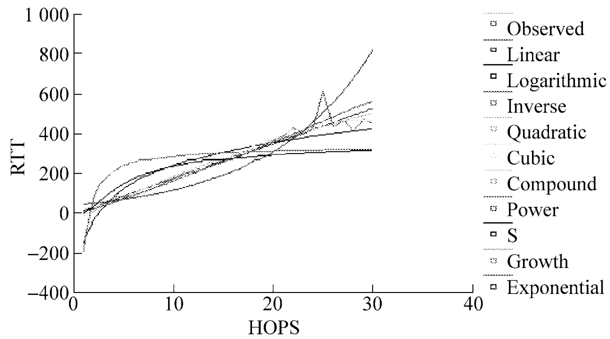

Fig. 6 The regression of hops and RTT

Figure 6 is the fitting result of 10 curves of Fig. 5. From the figure, it is obvious that liner, quadratic, cubic and power all approach the actual curve. Further analysis and comparison are shown as follows:

Table 4 Curve estimation

Mth	Rsq	d.f.	F	Sigf
LIN	0.934	28	393.91	0.000
LOG	0.815	28	123.08	0.000
INV	0.410	28	19.49	0.000
QUA	0.939	27	208.85	0.000
CUB	0.952	26	173.57	0.000
COM	0.765	28	91.03	0.000
POW	0.976	28	1 138.60	0.000
S	0.845	28	152.73	0.000
GRO	0.765	28	91.03	0.000
EXP	0.765	28	91.03	0.000

From the result of curve fitting, all the 10 models have statistical meaning, and the probability ratio of the values of

F variance analyzsis (Sigf) are small, near nearly 0.000. Compared to the value of adjusting (R_2), the R_2 value of power is the biggest, which makes the power model the best one with parameters of $b_0 = 11.617\,9$, $b_1 = 1.143\,7$.

It is discovered that the increase of RTT according to $Y = 11.617\,9X^{1.143\,7}$, is along with the increase of hops. The relationship of transmission time and nodes by which it passes takes on the power-law. In terms of separation degree transmission model Eq. (15), mean data transmission time of IP level is 116 millisecond during 2008 Beijing Olympic Games.

4 Conclusions

Heterogeneity, activity, decentralization, huge scale and complex topology of the Internet impel people using new theory and complex network theory to analyze it.

In this paper, we first considered the complex network theory and particularity of the Internet. The definition of the Internet separation degree was established and the mathematical relation between Internet separation degree and hops was founded. Secondly, in terms of CAIDA's data that were obtained from 30 global central routers in 4 years, the time series stochastic process madel of Internet separation degree was established, and its parameter was ascertained. Thirdly, correlation analysis of hops and RTT was done, and Internet separation degree time sensitivity model (ISDTSM) was established. It is proved that the transmission of the complex network of the Internet takes on the power-law ($Y = 11.617\,9X^{1.143\,7}$). Finally, the Internet separation degree and the mean time of data transmitting in IP level were forecasted respectively.

References

1. Floyd S., Paxson V., Difficulties in simulating the Internet., IEEE/ACM Trans. on Networking, 2001, 9(4): 392–403
2. Watts D., Strogatz S., Collective dynamics of 'small-world' networks, Nature, 1998, 393(6684): 440–442
3. Yook S. H., Jeong H., Barabasi A. L., Modeling the Internet's Large-scale Topology, PNAS 2002, 99: 13382–13386
4. Adamic L. A., Huberman B. A., Nature, 1999, 401: 131
5. AlbertR., Barabasi A. L., Rev.Mod.Phys., 2002
6. Faloutsos M.，Faloutsos P.，Faloutsos C., Comput, Commun, Rev, 1999, 29: 251
7. Govindan R., Tangmunarunkit H., Proceedings of IEEE INFOCOM 2000, Israel: Tel Aviv (IEEE, Piscataway, N.J.) 2000, 3: 1371
8. N. T. J. Bailey, The mathematical theory of infectious diseases and its applications, New York: Hafner Press, 1975
9. Anderson R. M., May R. M. Infectious diseases of humans, Oxford: Oxford University Press, 1992
10. Hethcote H. W, SIAM Review, 2000, 42, 599
11. Erdos P., Renyi A., 1959 Publ. Mat